

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Bujard, Hermann
Gossen, Manfred

(ii) TITLE OF INVENTION: Methods for Regulating Gene Expression

(iii) NUMBER OF SEQUENCES: 28

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: LAHIVE & COCKFIELD
(B) STREET: 60 State Street, suite 510
(C) CITY: Boston
(D) STATE: Massachusetts
(E) COUNTRY: USA
(F) ZIP: 02109-1875

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: ASCII Text

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA

(A) APPLICATION NUMBER: US 08/383,754
(B) FILING DATE: 03-FEB-1995
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA

(A) APPLICATION NUMBER: US 08/275,876
(B) FILING DATE: 15-JULY-94
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA

(A) APPLICATION NUMBER: US 08/270,637
(B) FILING DATE: 01-JULY-94
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA

(A) APPLICATION NUMBER: US 08/260,452
(B) FILING DATE: 14-JUNE-94
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA

(A) APPLICATION NUMBER: US 08/076,327
(B) FILING DATE: 14-JUNE-93
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA

(A) APPLICATION NUMBER: US 08/076,726

(B) FILING DATE: 14-JUNE-93

(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: DeConti, Giulio A. Jr.
(B) REGISTRATION NUMBER: 31,503
(C) REFERENCE/DOCKET NUMBER: BBI-009CP6

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (617) 227-7400
(B) TELEFAX: (617) 227-5941

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1008 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: exon
(B) LOCATION: 1..1008

(ix) FEATURE:

- (A) NAME/KEY: mRNA
(B) LOCATION: 1..1008

(ix) FEATURE:

- (A) NAME/KEY: misc. binding
(B) LOCATION: 1..207

(ix) FEATURE:

- (A) NAME/KEY: misc. binding
(B) LOCATION: 208..335

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 1..1005

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

45	ATG TCT AGA TTA GAT AAA AGT AAA GTG ATT AAC AGC GCA TTA GAG CTG	48
	Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu	
	1 5 10 15	
50	CTT AAT GAG GTC GGA ATC GAA GGT TTA ACA ACC CGT AAA CTC GCC CAG	96
	Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln	
	20 25 30	
55	AAG CTA GGT GTA GAG CAG CCT ACA CTG TAT TGG CAT GTA AAA AAT AAG	144
	Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys	
	35 40 45	

	CGG GCT TTG CTC GAC GCC TTA GCC ATT GAG ATG TTA GAT AGG CAC CAT	192
	Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His	
	50 55 60	
5	ACT CAC TTT TGC CCT TTA AAA GGG GAA AGC TGG CAA GAT TTT TTA CGC	240
	Thr His Phe Cys Pro Leu Lys Gly Glu Ser Trp Gln Asp Phe Leu Arg	
	65 70 75 80	
10	AAT AAG GCT AAA AGT TTT AGA TGT GCT TTA CTA AGT CAT CGC AAT GGA	288
	Asn Lys Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asn Gly	
	85 90 95	
15	GCA AAA GTA CAT TCA GAT ACA CGG CCT ACA GAA AAA CAG TAT GAA ACT	336
	Ala Lys Val His Ser Asp Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr	
	100 105 110	
20	CTC GAA AAT CAA TTA GCC TTT TTA TGC CAA CAA GGT TTT TCA CTA GAG	384
	Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu	
	115 120 125	
25	AAT GCA TTA TAT GCA CTC AGC GCT GTG GGG CAT TTT ACT TTA GGT TGC	432
	Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys	
	130 135 140	
30	GTA TTG GAA GAT CAA GAG CAT CAA GTC GCT AAA GAA GAA AGG GAA ACA	480
	Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr	
	145 150 155 160	
35	CCT ACT ACT GAT AGT ATG CCG CCA TTA TTA CGA CAA GCT ATC GAA TTA	528
	Pro Thr Thr Asp Ser Met Pro Pro Leu Arg Gln Ala Ile Glu Leu	
	165 170 175	
40	TTT GAT CAC CAA GGT GCA GAG CCA GCC TTC TTA TTC GGC CTT GAA TTG	576
	Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu	
	180 185 190	
45	ATC ATA TGC GGA TTA GAA AAA CAA CTT AAA TGT GAA AGT GGG TCC GCG	624
	Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser Ala	
	195 200 205	
50	TAC AGC CGC GCG CGT ACG AAA AAC AAT TAC GGG TCT ACC ATC GAG GGC	672
	Tyr Ser Arg Ala Arg Thr Lys Asn Asn Tyr Gly Ser Thr Ile Glu Gly	
	210 215 220	
55	CTG CTC GAT CTC CCG GAC GAC GAC GCC CCC GAA GAG GCG GGG CTG GCG	720
	Leu Leu Asp Leu Pro Asp Asp Asp Ala Pro Glu Glu Ala Gly Leu Ala	
	225 230 235 240	
60	GCT CCG CGC CTG TCC TTT CTC CCC GCG GGA CAC ACG CGC AGA CTG TCG	768
	Ala Pro Arg Leu Ser Phe Leu Pro Ala Gly His Thr Arg Arg Leu Ser	
	245 250 255	
65	ACG GCC CCC CCG ACC GAT GTC AGC CTG GGG GAC GAG CTC CAC TTA GAC	816
	Thr Ala Pro Pro Thr Asp Val Ser Leu Gly Asp Glu Leu His Leu Asp	
	260 265 270	
70	GGC GAG GAC GTG GCG ATG GCG CAT GCC GAC GCG CTA GAC GAT TTC GAT	864
	Gly Glu Asp Val Ala Met Ala His Ala Asp Ala Leu Asp Asp Phe Asp	
	275 280 285	

CTG GAC ATG TTG GGG GAC GGG GAT TCC CCG GGT CCG GGA TTT ACC CCC 912
Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly Pro Gly Phe Thr Pro
290 295 300

CAC GAC TCC GCC CCC TAC GGC GCT CTG GAT ATG GCC GAC TTC GAG TTT 960
His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met Ala Asp Phe Glu Phe
305 310 315 320

GAG CAG ATG TTT ACC GAT CCC CTT GGA ATT GAC GAG TAC GGT GGG TAG 1008
Glu Gln Met Phe Thr Asp Pro Leu Gly Ile Asp Glu Tyr Gly Gly
325 330 335

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 335 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu
1 5 10 15

Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln
20 25 30

Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys
35 40 45

Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His
50 55 60

Thr His Phe Cys Pro Leu Lys Gly Glu Ser Trp Gln Asp Phe Leu Arg
65 70 75 80

Asn Lys Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asn Gly
85 90 95

Ala Lys Val His Ser Asp Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr
100 105 110

Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu
115 120 125

Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys
130 135 140

Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr
145 150 155 160

Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu
165 170 175

Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu
180 185 190

Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser Ala
195 200 205

5 Tyr Ser Arg Ala Arg Thr Lys Asn Asn Tyr Gly Ser Thr Ile Glu Gly
210 215 220

Leu Leu Asp Leu Pro Asp Asp Asp Ala Pro Glu Glu Ala Gly Leu Ala
225 230 235 240

10 Ala Pro Arg Leu Ser Phe Leu Pro Ala Gly His Thr Arg Arg Leu Ser
245 250 255

15 Thr Ala Pro Pro Thr Asp Val Ser Leu Gly Asp Glu Leu His Leu Asp
260 265 270

Gly Glu Asp Val Ala Met Ala His Ala Asp Ala Leu Asp Asp Phe Asp
275 280 285

20 Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly Pro Gly Phe Thr Pro
290 295 300

His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met Ala Asp Phe Glu Phe
305 310 315 320

Glu Gln Met Phe Thr Asp Pro Leu Gly Ile Asp Glu Tyr Gly Gly
325 330 335

30 INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAC GCG CTA GAC GAT TTC GAT CTG GAC ATG TTG 33
Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
1 5 10

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
1 5 10

10 (2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
15 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal
20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

25 Met Pro Lys Arg Pro Arg Pro
1 5

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 569 base pairs
(B) TYPE: nucleic acid
30 (C) STRANDEDNESS: double
(D) TOPOLOGY: linear
35

(ii) MOLECULE TYPE: DNA

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GAATTCGGGG CCGCGGAGGC TGGATCGGTC CCGGTGTCTT CTATGGAGGT CAAAACAGCG	60
TGGATGGCGT CTCCAGGCGA TCTGACGGTT CACTAAACGA GCTCTGCTTA TATAGGTCGA	120
45 GTTTACCACT CCCTATCAGT GATAGAGAAA AGTGAAAGTC GAGTTTACCA CTCCCTATCA	180
GTGATAGAGA AAAGTGAAAG TCGAGTTTAC CACTCCCTAT CAGTGATAGA GAAAAGTGAA	240
50 AGTCGAGTTT ACCACTCCCT ACCAGTGATA GAGAAAAGTG AAAGTCGAGT TTACCACTCC	300
CTATCAGTGA TAGAGAAAAG TGAAAGTCGA GTTTACCACT CCCTATCAGT GATAGAGAAA	360
AGTGAAAGTC GAGTTTACCA CTCCCTATCA GTGATAGAGA AAAGTGAAAG TCGAGCTCGG	420
55 TACCCGGGTC GAGTAGGCGT GTACGGTGGG AGGCCTATAT AAGCAGAGCT CGTTTAGTGA	480
ACCGTCAGAT CGCCTGGAGA CGCCATCCAC GCTGTTTTGA CCTCCATAGA AGACACCGGG	540

ACCGATCCAG CCTCCGCGGC CCCGAATTC

569

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 520 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AGATCTGCAG	GGTCGCTCGG	TGTTTCGAGGC	CACACGCGTC	ACCTTAATAT	GCGAAGTGGA	60
CCGGATCTCG	AGTTTACCAC	TCCCTATCAG	TGATAGAGAA	AAGTGAAAGT	CGAGTTTACC	120
ACTCCCTATC	AGTGATAGAG	AAAAGTGAAA	GTCGAGTTTA	CCACTCCCTA	TCAGTGATAG	180
AGAAAAGTGA	AAGTCGAGTT	TACCACTCCC	TATCAGTGAT	AGAGAAAAGT	GAAAGTCGAG	240
TTTACCACTC	CCTATCAGTG	ATAGAGAAAA	GTGAAAGTCG	AGTTTACCAC	TCCCTATCAG	300
TGATAGAGAA	AAGTGAAAGT	CGAGTTTACC	ACTCCCTATC	AGTGATAGAG	AAAAGTGAAA	360
GTCGAGCTCG	GTACCCGGGT	CGAGTAGGCG	TGTACGGTGG	GAGGCCTATA	TAAGCAGAGC	420
TCGTTTAGTG	AACCGTCAGA	TCGCCTGGAG	ACGCCATCCA	CGCTGTTTTG	ACCTCCATAG	480
AAGACACCGG	GACCGATCCA	GCCTCCGCGG	CCCCGAATTC			520

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 450 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Human cytomegalovirus
- (B) STRAIN: K12, Towne

(ix) FEATURE:

- (A) NAME/KEY: mRNA
- (B) LOCATION: 382..450

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAATTCCTCG	AGTTTACCAC	TCCCTATCAG	TGATAGAGAA	AAGTGAAAGT	CGAGTTTACC	60
ACTCCCTATC	AGTGATAGAG	AAAAGTGAAA	GTCGAGTTTA	CCACTCCCTA	TCAGTGATAG	120

AGAAAAGTGA AAGTCGAGTT TACCACTCCC TATCAGTGAT AGAGAAAAGT GAAAGTCGAG 180
TTTACCACTC CCTATCAGTG ATAGAGAAAA GTGAAAGTCG AGTTTACCAC TCCCTATCAG 240
TGATAGAGAA AAGTGAAAGT CGAGTTTACC ACTCCCTATC AGTGATAGAG AAAAGTGAAA 300
GTCGAGCTCG GTACCCGGGT CGAGTAGGCG TGTACGGTGG GAGGCCTATA TAAGCAGAGC 360
5 TCGTTTAGTG AACCGTCAGA TCGCCTGGAG ACGCCATCCA CGCTGTTTTG ACCTCCATAG 420
AAGACACCGG GACCGATCCA GCCTCCGCGG 450

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 450 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Human cytomegalovirus
(B) STRAIN: Towne

(ix) FEATURE:

- (A) NAME/KEY: mRNA
(B) LOCATION: 382..450

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GAATTCCTCG ACCCGGGTAC CGAGCTCGAC TTCACTTTT CTCTATCACT GATAGGGAGT 60
GGTAAACTCG ACTTTCACCT TTCTCTATCA CTGATAGGGA GTGGTAAACT CGACTTTCAC 120
TTTTCTCTAT CACTGATAGG GAGTGGTAAA CTCGACTTTC ACTTTTCTCT ATCACTGATA 180
GGGAGTGGTA AACTCGACTT TCACTTTTCT CTATCACTGA TAGGGAGTGG TAAACTCGAC 240
30 TTTCACTTTT CTCTATCACT GATAGGGAGT GGTAAACTCG ACTTTCACCT TTCTCTATCA 300
CTGATAGGGA GTGGTAAACT CGAGTAGGCG TGTACGGTGG GAGGCCTATA TAAGCAGAGC 360
TCGTTTAGTG AACCGTCAGA TCGCCTGGAG ACGCCATCCA CGCTGTTTTG ACCTCCATAG 420
AAGACACCGG GACCGATCCA GCCTCCGCGG 450

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 398 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Herpes Simplex Virus

(B) STRAIN: KOS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GAGCTCGACT TTCACTTTTC TCTATCACTG ATAGGGAGTG GTAAACTCGA CTTTCACTTT 60
TCTCTATCAC TGATAGGGAG TGGTAAACTC GACTTTCACT TTTCTCTATC ACTGATAGGG 120
AGTGGTAAAC TCGACTTTCA CTTTCTCTA TCACTGATAG GGAGTGGTAA ACTCGACTTT 180
CACTTTTCTC TATCACTGAT AGGGAGTGGT AAACTCGACT TTCACTTTTC TCTATCACTG 240
ATAGGGAGTG GTAAACTCGA CTTTCACTTT TCTCTATCAC TGATAGGGAG TGGTAAACTC 300
GAGATCCGGC GAATTCTGAAC ACGCAGATGC AGTCGGGGCG GCGCGGTCCG AGGTCCACTT 360
CGCATATTAA GGTGACGCGT GTGGCCTCGA ACACCGAG 398

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACTTTATCAC TGATAAACAA ACTTATCAGT GATAAAGA 38

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ACTCTATCAT TGATAGAGTT CCCTATCAGT GATAGAGA 38

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGCTTATCAT CGATAAGCTA GTTTATCACA GTTAAATT

38

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ACTCTATCAT TGATAGGGAA CTCTATCAAT GATAGGGA

38

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AATCTATCAC TGATAGAGTA CCCTATCATC GATAGAGA

38

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 621 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

5	ATG TCT AGA TTA GAT AAA AGT AAA GTG ATT AAC AGC GCA TTA GAG CTG	48
	Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu	
	1 5 10 15	
10	CTT AAT GAG GTC GGA ATC GAA GGT TTA ACA ACC CGT AAA CTC GCC CAG	96
	Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln	
	20 25 30	
15	AAG CTA GGT GTA GAG CAG CCT ACA TTG TAT TGG CAT GTA AAA AAT AAG	144
	Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys	
	35 40 45	
20	CGG GCT TTG CTC GAC GCC TTA GCC ATT GAG ATG TTA GAT AGG CAC CAT	192
	Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His	
	50 55 60	
25	ACT CAC TTT TGC CCT TTA GAA GGG GAA AGC TGG CAA GAT TTT TTA CGT	240
	Thr His Phe Cys Pro Leu Glu Gly Glu Ser Trp Gln Asp Phe Leu Arg	
	65 70 75 80	
30	AAT AAG GCT AAA AGT TTT AGA TGT GCT TTA CTA AGT CAT CGC GAT GGA	288
	Asn Lys Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asp Gly	
	85 90 95	
35	GCA AAA GTA CAT TTA GGT ACA CGG CCT ACA GAA AAA CAG TAT GAA ACT	336
	Ala Lys Val His Leu Gly Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr	
	100 105 110	
40	CTC GAA AAT CAA TTA GCC TTT TTA TGC CAA CAA GGT TTT TCA CTA GAG	384
	Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu	
	115 120 125	
45	AAT GCA TTA TAT GCA CTC AGC GCT GTG GGG CAT TTT ACT TTA GGT TGC	432
	Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys	
	130 135 140	
50	GTA TTG GAA GAT CAA GAG CAT CAA GTC GCT AAA GAA GAA AGG GAA ACA	480
	Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr	
	145 150 155 160	
55	CCT ACT ACT GAT AGT ATG CCG CCA TTA TTA CGA CAA GCT ATC GAA TTA	528
	Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu	
	165 170 175	
60	TTT GAT CAC CAA GGT GCA GAG CCA GCC TTC TTA TTC GGC CTT GAA TTG	576
	Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu	
	180 185 190	
65	ATC ATA TGC GGA TTA GAA AAA CAA CTT AAA TGT GAA AGT GGG TCC	621
	Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser	
	195 200 205	

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 207 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

	Met	Ser	Arg	Leu	Asp	Lys	Ser	Lys	Val	Ile	Asn	Ser	Ala	Leu	Glu	Leu	
	1				5					10					15		
10	Leu	Asn	Glu	Val	Gly	Ile	Glu	Gly	Leu	Thr	Thr	Arg	Lys	Leu	Ala	Gln	
				20					25					30			
	Lys	Leu	Gly	Val	Glu	Gln	Pro	Thr	Leu	Tyr	Trp	His	Val	Lys	Asn	Lys	
			35					40					45				
15	Arg	Ala	Leu	Leu	Asp	Ala	Leu	Ala	Ile	Glu	Met	Leu	Asp	Arg	His	His	
		50					55					60					
	Thr	His	Phe	Cys	Pro	Leu	Glu	Gly	Glu	Ser	Trp	Gln	Asp	Phe	Leu	Arg	
20		65				70					75					80	
	Asn	Lys	Ala	Lys	Ser	Phe	Arg	Cys	Ala	Leu	Leu	Ser	His	Arg	Asp	Gly	
					85					90					95		
25	Ala	Lys	Val	His	Leu	Gly	Thr	Arg	Pro	Thr	Glu	Lys	Gln	Tyr	Glu	Thr	
				100					105					110			
	Leu	Glu	Asn	Gln	Leu	Ala	Phe	Leu	Cys	Gln	Gln	Gly	Phe	Ser	Leu	Glu	
			115				120						125				
30	Asn	Ala	Leu	Tyr	Ala	Leu	Ser	Ala	Val	Gly	His	Phe	Thr	Leu	Gly	Cys	
		130				135						140					
	Val	Leu	Glu	Asp	Gln	Glu	His	Gln	Val	Ala	Lys	Glu	Glu	Arg	Glu	Thr	
35		145				150				155					160		
	Pro	Thr	Thr	Asp	Ser	Met	Pro	Pro	Leu	Leu	Arg	Gln	Ala	Ile	Glu	Leu	
					165					170					175		
40	Phe	Asp	His	Gln	Gly	Ala	Glu	Pro	Ala	Phe	Leu	Phe	Gly	Leu	Glu	Leu	
				180					185					190			
	Ile	Ile	Cys	Gly	Leu	Glu	Lys	Gln	Leu	Lys	Cys	Glu	Ser	Gly	Ser		
45			195				200					205					

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 621 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

5	ATG TCT AGA TTA GAT AAA AGT AAA GTG ATT AAC AGC GCA TTA GAG CTG	48
	Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu	
	1 5 10 15	
10	CTT AAT GAG GTC GGA ATC GAA GGT TTA ACA ACC CGT AAA CTC GCC CAG	96
	Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln	
	20 25 30	
15	AAG CTA GGT GTA GAG CAG CCT ACA CTG TAT TGG CAT GTA AAA AAT AAG	144
	Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys	
	35 40 45	
20	CGG GCT TTG CTC GAC GCC TTA GCC ATT GAG ATG TTA GAT AGG CAC CAT	192
	Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His	
	50 55 60	
25	ACT CAC TTT TGC CCT TTA AAA GGG GAA AGC TGG CAA GAT TTT TTA CGC	240
	Thr His Phe Cys Pro Leu Lys Gly Glu Ser Trp Gln Asp Phe Leu Arg	
	65 70 75 80	
30	AAT AAG GCT AAA AGT TTT AGA TGT GCT TTA CTA AGT CAT CGC AAT GGA	288
	Asn Lys Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asn Gly	
	85 90 95	
35	GCA AAA GTA CAT TCA GAT ACA CGG CCT ACA GAA AAA CAG TAT GAA ACT	336
	Ala Lys Val His Ser Asp Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr	
	100 105 110	
40	CTC GAA AAT CAA TTA GCC TTT TTA TGC CAA CAA GGT TTT TCA CTA GAG	384
	Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu	
	115 120 125	
45	AAT GCA TTA TAT GCA CTC AGC GCT GTG GGG CAT TTT ACT TTA GGT TGC	432
	Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys	
	130 135 140	
50	GTA TTG GAA GAT CAA GAG CAT CAA GTC GCT AAA GAA GAA AGG GAA ACA	480
	Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr	
	145 150 155 160	
55	CCT ACT ACT GAT AGT ATG CCG CCA TTA TTA CGA CAA GCT ATC GAA TTA	528
	Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu	
	165 170 175	
60	TTT GAT CAC CAA GGT GCA GAG CCA GCC TTC TTA TTC GGC CTT GAA TTG	576
	Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu	
	180 185 190	
65	ATC ATA TGC GGA TTA GAA AAA CAA CTT AAA TGT GAA AGT GGG TCC	621
	Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser	
	195 200 205	

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 207 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```

Met Ser Arg Leu Asp Lys Ser Lys Val Ile Asn Ser Ala Leu Glu Leu
 1           5           10           15
10 Leu Asn Glu Val Gly Ile Glu Gly Leu Thr Thr Arg Lys Leu Ala Gln
    20           25           30
    Lys Leu Gly Val Glu Gln Pro Thr Leu Tyr Trp His Val Lys Asn Lys
    35           40           45
15 Arg Ala Leu Leu Asp Ala Leu Ala Ile Glu Met Leu Asp Arg His His
    50           55           60
    Thr His Phe Cys Pro Leu Lys Gly Glu Ser Trp Gln Asp Phe Leu Arg
20 65           70           75           80
    Asn Lys Ala Lys Ser Phe Arg Cys Ala Leu Leu Ser His Arg Asn Gly
    85           90           95
25 Ala Lys Val His Ser Asp Thr Arg Pro Thr Glu Lys Gln Tyr Glu Thr
    100          105          110
    Leu Glu Asn Gln Leu Ala Phe Leu Cys Gln Gln Gly Phe Ser Leu Glu
    115          120          125
30 Asn Ala Leu Tyr Ala Leu Ser Ala Val Gly His Phe Thr Leu Gly Cys
    130          135          140
    Val Leu Glu Asp Gln Glu His Gln Val Ala Lys Glu Glu Arg Glu Thr
35 145          150          155          160
    Pro Thr Thr Asp Ser Met Pro Pro Leu Leu Arg Gln Ala Ile Glu Leu
    165          170          175
40 Phe Asp His Gln Gly Ala Glu Pro Ala Phe Leu Phe Gly Leu Glu Leu
    180          185          190
    Ile Ile Cys Gly Leu Glu Lys Gln Leu Lys Cys Glu Ser Gly Ser
45 195          200          205

```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 192 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

	GAC	ATG	GAA	AAA	GCG	ACA	CCG	GAG	ACG	ATG	GTC	CAT	TGG	ATT	TGT	CTG	48
	Asp	Met	Glu	Lys	Ala	Thr	Pro	Glu	Thr	Met	Val	His	Trp	Ile	Cys	Leu	
5	1				5					10					15		
	AAG	ATG	GAG	CCA	GCT	CTG	TGG	ATG	GCC	ATT	ACA	GCA	ACA	TCG	CAC	GGC	96
	Lys	Met	Glu	Pro	Ala	Leu	Trp	Met	Ala	Ile	Thr	Ala	Thr	Ser	His	Gly	
				20					25					30			
10	GCA	AGG	CAC	AGG	ACA	TTC	GTC	GGG	TTT	TCC	GGC	TGC	CTC	CAC	CGC	AAA	144
	Ala	Arg	His	Arg	Thr	Phe	Val	Gly	Phe	Ser	Gly	Cys	Leu	His	Arg	Lys	
			35					40					45				
15	TCC	CTC	ACG	TAC	CCA	GTG	ATA	TGC	CTG	AGC	AAA	CCG	AGC	CAG	AGG	ATT	192
	Ser	Leu	Thr	Tyr	Pro	Val	Ile	Cys	Leu	Ser	Lys	Pro	Ser	Gln	Arg	Ile	
		50					55					60					

20

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 64 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

	Asp	Met	Glu	Lys	Ala	Thr	Pro	Glu	Thr	Met	Val	His	Trp	Ile	Cys	Leu
30	1				5					10					15	
	Lys	Met	Glu	Pro	Ala	Leu	Trp	Met	Ala	Ile	Thr	Ala	Thr	Ser	His	Gly
				20					25					30		
35	Ala	Arg	His	Arg	Thr	Phe	Val	Gly	Phe	Ser	Gly	Cys	Leu	His	Arg	Lys
			35					40					45			
	Ser	Leu	Thr	Tyr	Pro	Val	Ile	Cys	Leu	Ser	Lys	Pro	Ser	Gln	Arg	Ile
		50					55					60				

40

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 816 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

45

(ii) MOLECULE TYPE: cDNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

	CTG	GAC	GAC	TCG	AAG	CGC	GTA	GCC	AAG	CGG	AAG	CTG	ATC	GAG	GAG	AAC	48
	Leu	Asp	Asp	Ser	Lys	Arg	Val	Ala	Lys	Arg	Lys	Leu	Ile	Glu	Glu	Asn	
	1				5				10					15			
5	CGG	GAG	CGG	CGA	CGC	AAG	GAG	GAG	ATG	ATC	AAA	TCC	CTG	CAG	CAC	CGG	96
	Arg	Glu	Arg	Arg	Arg	Lys	Glu	Glu	Met	Ile	Lys	Ser	Leu	Gln	His	Arg	
				20					25					30			
10	CCC	AGC	CCC	AGC	GCA	GAG	GAG	TGG	GAG	CTG	ATC	CAC	GTG	GTG	ACC	GAG	144
	Pro	Ser	Pro	Ser	Ala	Glu	Glu	Trp	Glu	Leu	Ile	His	Val	Val	Thr	Glu	
			35					40					45				
15	GCG	CAC	CGC	AGC	ACC	AAC	GCG	CAG	GGC	AGC	CAC	TGG	AAG	CAG	AGG	AGG	192
	Ala	His	Arg	Ser	Thr	Asn	Ala	Gln	Gly	Ser	His	Trp	Lys	Gln	Arg	Arg	
		50					55					60					
20	AAA	TTC	CTG	CTC	GAA	GAT	ATC	GGT	CAG	TCG	CCC	ATG	GCC	TCC	ATG	CTT	240
	Lys	Phe	Leu	Leu	Glu	Asp	Ile	Gly	Gln	Ser	Pro	Met	Ala	Ser	Met	Leu	
	65					70					75					80	
25	GAC	GGG	GAC	AAA	GTG	GAC	CTG	GAG	GCG	TTC	AGC	GAG	TTT	ACA	AAA	ATC	288
	Asp	Gly	Asp	Lys	Val	Asp	Leu	Glu	Ala	Phe	Ser	Glu	Phe	Thr	Lys	Ile	
					85					90					95		
30	ATC	ACG	CCG	GCC	ATC	ACC	CGC	GTG	GTC	GAC	TTT	GCC	AAA	AAC	CTG	CCC	336
	Ile	Thr	Pro	Ala	Ile	Thr	Arg	Val	Val	Asp	Phe	Ala	Lys	Asn	Leu	Pro	
				100					105					110			
35	ATG	TTC	TCG	GAG	CTG	CCG	TGC	GAG	GAT	CAG	ATC	ATC	CTG	CTG	AAG	GGC	384
	Met	Phe	Ser	Glu	Leu	Pro	Cys	Glu	Asp	Gln	Ile	Ile	Leu	Leu	Lys	Gly	
				115				120					125				
40	TGC	TGC	ATG	GAG	ATC	ATG	TCG	CTG	CGC	GCC	GCC	GTG	CGC	TAC	GAC	CCC	432
	Cys	Cys	Met	Glu	Ile	Met	Ser	Leu	Arg	Ala	Ala	Val	Arg	Tyr	Asp	Pro	
		130					135					140					
45	GAG	AGC	GAA	ACG	CTG	ACG	CTG	AGC	GGG	GAA	ATG	GCC	GTC	AAA	CGC	GAG	480
	Glu	Ser	Glu	Thr	Leu	Thr	Leu	Ser	Gly	Glu	Met	Ala	Val	Lys	Arg	Glu	
	145					150					155					160	
50	CAG	TTG	AAG	AAC	GGA	GGG	CTG	GGG	GTC	GTG	TCT	GAT	GCC	ATC	TTC	GAC	528
	Gln	Leu	Lys	Asn	Gly	Gly	Leu	Gly	Val	Val	Ser	Asp	Ala	Ile	Phe	Asp	
				165					170						175		
55	CTC	GGC	AAG	TCG	CTG	TCT	GCC	TTC	AAC	CTG	GAC	GAC	ACC	GAG	GTG	GCC	576
	Leu	Gly	Lys	Ser	Leu	Ser	Ala	Phe	Asn	Leu	Asp	Asp	Thr	Glu	Val	Ala	
				180					185					190			
60	CTG	CTG	CAG	GCC	GTG	CTG	CTC	ATG	TCC	TCA	GAC	CGG	ACG	GGG	CTG	ATC	624
	Leu	Leu	Gln	Ala	Val	Leu	Leu	Met	Ser	Ser	Asp	Arg	Thr	Gly	Leu	Ile	
				195				200					205				
65	TGC	GTG	GAT	AAG	ATA	GAG	AAG	TGC	CAG	GAG	TCG	TAC	CTG	CTG	GCG	TTC	672
	Cys	Val	Asp	Lys	Ile	Glu	Lys	Cys	Gln	Glu	Ser	Tyr	Leu	Leu	Ala	Phe	
		210					215					220					
70	GAG	CAC	TAC	ATC	AAC	TAC	CGC	AAA	CAC	AAC	ATT	CCC	CAC	TTC	TGG	TCC	720
	Glu	His	Tyr	Ile	Asn	Tyr	Arg	Lys	His	Asn	Ile	Pro	His	Phe	Trp	Ser	
	225					230					235					240	

AAG CTG CTG ATG AAG GTG GCG GAC CTG CGC ATG ATC GGC GCC TAC CAC 768
Lys Leu Leu Met Lys Val Ala Asp Leu Arg Met Ile Gly Ala Tyr His
245 250 255

GCC AGC CGC TTC CTG CAC ATG AAG GTG GAG TGC CCC ACC GAG CTC TCC 816
Ala Ser Arg Phe Leu His Met Lys Val Glu Cys Pro Thr Glu Leu Ser
260 265 270

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 272 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Leu Asp Asp Ser Lys Arg Val Ala Lys Arg Lys Leu Ile Glu Glu Asn
1 5 10 15

Arg Glu Arg Arg Arg Lys Glu Glu Met Ile Lys Ser Leu Gln His Arg
20 25 30

Pro Ser Pro Ser Ala Glu Glu Trp Glu Leu Ile His Val Val Thr Glu
35 40 45

Ala His Arg Ser Thr Asn Ala Gln Gly Ser His Trp Lys Gln Arg Arg
50 55 60

Lys Phe Leu Leu Glu Asp Ile Gly Gln Ser Pro Met Ala Ser Met Leu
65 70 75 80

Asp Gly Asp Lys Val Asp Leu Glu Ala Phe Ser Glu Phe Thr Lys Ile
85 90 95

Ile Thr Pro Ala Ile Thr Arg Val Val Asp Phe Ala Lys Asn Leu Pro
100 105 110

Met Phe Ser Glu Leu Pro Cys Glu Asp Gln Ile Ile Leu Leu Lys Gly
115 120 125

Cys Cys Met Glu Ile Met Ser Leu Arg Ala Ala Val Arg Tyr Asp Pro
130 135 140

Glu Ser Glu Thr Leu Thr Leu Ser Gly Glu Met Ala Val Lys Arg Glu
145 150 155 160

Gln Leu Lys Asn Gly Gly Leu Gly Val Val Ser Asp Ala Ile Phe Asp
165 170 175

Leu Gly Lys Ser Leu Ser Ala Phe Asn Leu Asp Asp Thr Glu Val Ala
180 185 190

Leu Leu Gln Ala Val Leu Leu Met Ser Ser Asp Arg Thr Gly Leu Ile

195 200 205

Cys Val Asp Lys Ile Glu Lys Cys Gln Glu Ser Tyr Leu Leu Ala Phe
210 215 220

Glu His Tyr Ile Asn Tyr Arg Lys His Asn Ile Pro His Phe Trp Ser
225 230 235 240

Lys Leu Leu Met Lys Val Ala Asp Leu Arg Met Ile Gly Ala Tyr His
245 250 255

Ala Ser Arg Phe Leu His Met Lys Val Glu Cys Pro Thr Glu Leu Ser
260 265 270

15 (2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24.

TCCCCGGGTA ACTAAGTAAG GATCC

25

25 (2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

35 AGTGGGTCCC CGGGTGACAT GGAA

24

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: polypeptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Ser Gly Ser Pro Gly Asp Met Glu

1

5

(2) INFORMATION FOR SEQ ID NO:27:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

AGTGGGTCCC CGGGTCTGGA CGAC

24

15

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: polypeptide

(v) FRAGMENT TYPE: internal

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Ser Gly Ser Pro Gly Leu Asp Asp

1

5